

【数据库建设】

ILASII 采访子系统及相关数据库的维护

洪 丹萍(厦门大学图书馆,福建 厦门 361005)

[摘 要] 随着计算机技术、网络技术与电子商务的迅速发展,大部分图书馆工作逐步实现了自动化管理。本文结合我校图书馆的实际工作,探讨如何快速、准确地将书店、出版商提供的电子版采访数据,利用图书馆自动化集成系统所自有的功能,通过数据接口将采访数据批接入相关的数据库;并探讨了 ILASII 采访子系统所涉及的相关数据库的数据维护与恢复等。

[关键词] ILASII 系统;采访数据;数据库维护

[中图分类号] G250.74 [文献标识码] B [文章编号] 1005-6610(2007)01-0062-04

随着计算机技术、网络技术的高速发展与图书馆自动化集成系统的不断完善,大部分图书馆工作逐步实现了自动化管理。高校图书馆采访部门订购图书的采访手段,也由现代化逐步替代了传统的图书采访模式,大大提高了工作效率。厦门大学图书馆目前使用的自动化集成系统是 ILASII2.0 版。ILAS II 是采用客户机/服务器(Client/Server)模式,主服务器采用 UNIX 平台,客户端以 WIN98、WIN2000 为依托。ILASII 提供了数据接收、数据输出接口以及 MARC 数据和定长数据的批处理等功能,为采编部门涉及的相关数据库的数据输出、数据接收以及数据的维护提供了方便。

1 采访子系统涉及的主要数据库

采访子系统是针对图书馆的采访业务设立的专用管理模块,实现对各种文献载体(包括普通图书、视听资料等)进入流通管理前的订购、验收、财经管理、统计及各种产品的输出等工作。采访子系统所涉及的数据库包括征订书目库、采编书目库、采访业务库、采访财经库、采编馆藏库、中央书目库、外部书目库等。

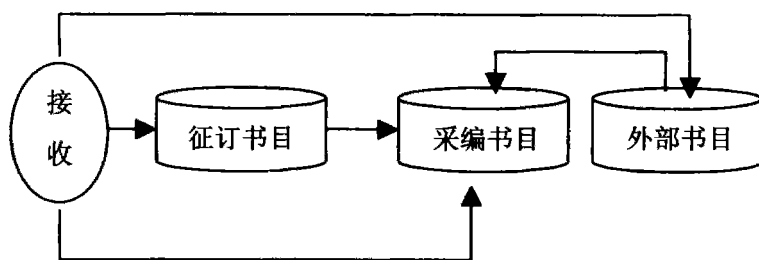
征订书目库采用标准 MARC 格式,图书馆需要进行订购处理的数据必须首先将书目数据接入或录入征订书目库。征订书目中除书目特征外,还包括征订目录名、征订目录的征订期(年)的信息和征订截止日期。采编书目库采用标准 MARC 格式。采访工作人员在“订购处理”界面,通过对本馆“分编库”和“总书目库”数据进行查重后,决定是否订购该图书,如从“征订书目列表”中确认订购某一条记录,即在采编书目库增加一条书目记录,所以采编书目库是存放采编工作中未经审校交送的书目数据。

采访业务库采用定长+变长格式,数据字段以 ACQ 为起始标识。采访业务库数据与书目数据密不可分,揭示各馆采访业务状况,主要分订购、验收(订购验收/直接验收)、退订三类记录。

2 图书馆电子版采访数据的批接收

随着电子商务的迅速发展,很多书店、出版社建立了自己的网站,提供网上服务方式,并能提供标准化的可操作的电子版书目数据。采访工作人员不用逐条将要订购的图书数据录入相关的数据库,而可利用图书馆自动化集成系统所自有的功能,通过数据接口将采访数据批接入相关的数据库,从而大大提高工作效率。

我馆采访部门对于书商提供的电子版书目数据一般采用以下 3 种方法接入 ILASII 数据库(如下图)。



2.1 电子版书目数据入外部书目库

如果书商提供的电子版数据是中文 MARC 格式, 可通过“MARC 数据”→“MARC 库”将数据接入中文外部书目库, (如果是西文 MARC 格式, 可接入西文外部书目库)。

①进入 ILASII“系统管理员”→“系统工具”→“MARC 数据”→“MARC 库”界面, ②“接收数据文本文件”通过“浏览”选择书商提供的电子版数据文件, ③“书目库选择”通过下拉菜单选择“外部书目库”, ④“书目记录按原号”一般不打勾, 系统可自动分配累加记录号, ⑤如果书商提供的数据没有控制号(没有 001 字段), “控制号按原号”可不打勾, 然后在“起始号”设置自己的控制号, 方便以后的数据维护, 如在“起始号”栏键入“XMU 0600001”(10 位数), ⑥为了保险起见, 最好先点击“第一条”、“下一条”, 查看数条数据, 确认正确后再点击“接收”。

由于采用这种方法是通过“系统管理员”界面进入, 将书商提供的电子版书目数据接入“外部书目库”, 在该功能界面, 可供选择的数据库较多, 所以要特别小心使用, 正确选择好接收数据的数据库。

2.2 电子版书目数据入采编书目库

也可以通过“编目管理”模块的“数据接口”, 将数据接入“采编书目库”。操作方法大致与接收数据入外部书目库相同。

①进入“编目管理”→“数据接口”→“接收 MARC 数据”界面, ②“接收数据文本文件”通过“浏览”选择书商提供的电子版数据文件, ③“书目库选择”通过下拉菜单选择“采编书目库”, ④~⑥同接收数据入外部书目库。

将数据接入采编书目库, 该操作可以从编目子系统的数据接口进去, 该接口仅提供“采编书目库”和“中央书目库”两个数据库的选择。采编书目库存放采编工作中未经审校交送的书目数据, 数据经审校后进入中央书目库。应该说, 采编书目库的数据是订购基础数据, 记录号要与对应的中央书目库取得一致。因此, 采用这种方法将数据接入“采编书目库”, 一般必须确定哪些数据全部, 或大部分是要订购的, 最终要进入中央书目库的, 假如接收进去的数据中含有无效数据, 将成为采编库的“垃圾”, 最终必须删除, 从而造成书目数据记录号不连续。

2.3 电子版书目数据入征订目录库

还可以通过“采访管理”→“征订目录管理”, 将数据接入“征订目录库”

①进入“采访管理”→“征订目录管理”界面, ②在总征订目录下添加分目录, ③选中新添加的分目录, 点击“接收”, ④在弹出的框中: “源数据文件”通过“浏览”选择书商提供的电子版数据文件, ⑤根据书商提供的电子版书目数据的格式, 选择“MARC 格式”或“定长数据”或“带分隔符定长数据”, 如果是定长数据或带分隔符定长数据要设置“定长格式文件”, “书目记录号”一般取“自动取号”, “书目控制号”选择“自动取号”, 然后设置“起始控制号”, 如“XMU 0000001”, ⑥填上“征订目录名”和“征订期(年)”, 然后点击“正式转换”。

该方法是通过采访子系统征订目录管理的数据接收功能将数据接入征订目录库, 征订目录是订购处理的核心, 也是网上采访事务的关键数据。征订目录包括征订目录名、征订期号、征订书目三层结构。采访人员通过征订目录, 经查重、确认书目数据, 然后进行订购; 也可接收来自网上的分馆或部门的报订

信息, 汇总后作为当前订购的分配。因此, 笔者认为, 通过“征订目录管理”将书商提供的电子版数据接入采访的“征订目录库”是最佳的方法。

3 采访子系统相关数据库的维护

3.1 征订目录库的维护

3.1.1 征订目录库数据的排序问题

采访工作人员将电子版数据接入“征订目录库”后, 接收进去的数据有时可以按征订序号排序, 但有时会出现杂乱无序, 不方便工作人员处理。为什么会出现这样的情况呢? 通过反复数次的测试, 发现要避免接收的数据不能排序的问题, 必须在接收数据时, 征订期(年)名称最好不要超过8个字节数, 如图(一), “征订期(年)”栏键入2006JL001, 超过8个字节数, 所接收的数据就会存在不能排序的问题, 如图(二)。

征订目录名	交流中心
征订期(年)	2006JL001

图(一)

征订序号	题名
000001	统计学
000004	专家指点·图解的技巧
000013	信息系统管理
000014	领导者之剑: 成功人士的5大突破思维技巧: t
000029	转型时期中国旅游产业环境、制度与模式研究
000030	中西文化比较概论
000031	中国文化年鉴: 2004
000044	谈判者心智
征订书目列表 详细征订书目 预订信息 荐购信息	

图(二)

3.1.2 征订目录库数据的维护

征订目录是采访工作人员订购处理的核心, 反反复复的数据接收、数据字段的批增、减处理、订购处理以及征订目录期的增、减等。对更新频繁、实时索引的数据库来说, 重组是一件不能回避的工作。重组不仅可以增减索引字段, 而且可以改善存储与检索效率。一般情况下对征订目录库数据的维护方法是“重组”数据库。

进入 database 目录, 键入命令: ../bin/ldb_rebuild ordcata. db

或者采用“重组→压缩→再重组”的方法来恢复数据库。

即: 进入 database 目录

(1) 重组数据库: ../bin/ldb_rebuild ordcata. db

(2) 压缩数据库: ../bin/ldb_compress ordcata. dat

(3) 再次重组数据库: cp LDBMS. TMP ordcata. dat

../bin/ldb_rebuild ordcata. db

(压缩数据库后会在 database 目录下生成 LDBMS. TMP 文件, 复制 LDBMS. TMP 文件后再次重组数据库)

但是, 在特殊情况下, 采用了“重组”或“重组→压缩→再重组”的方法都无法恢复数据库。

我馆曾出现工作人员新增征订目录后无法接收数据,要删除该征订目录出现“当前征订目录不能定位”的错误信息。采用了一般的恢复方法——“重组”数据库和“重组→压缩→再重组”都无济于事。为应付意外事件的发生,我馆每天都会将数据打包、压缩后传输到另一台服务器备用,因此,遇到这种情况,就可以取完好的备份数据来恢复。

操作步骤如下:

(1) 将经压缩、打包的备份数据传输到服务器。

(2) 解压、拆包备份数据

(3) 取出征订目录库数据 `ordcata.dat`、`ordcata.idx` 和 `ordcata.db` 覆盖到 `database` 目录下。如果条件允许,最好数据恢复后再重组。

3.2 外部书目库的维护

中文外部书目库(`middles`)存放着我馆从北图购买的书目数据以及工作人员从网上下载的书目数据等,数据库非常庞大,假如 `middles` 出问题,维护需要非常长的时间,必将影响图书馆的正常工作以及对读者的服务。因此,笔者利用放长假,闭馆时间长,将中文外部书目库做了彻底的处理。

(1) 先进入 ILASII“系统管理员”→“系统工具”→“MARC 数据批处理”界面,选择“外部书目库”,删除 `middles` 数据库中多余数据,即采访部门反复批接入中文外部书目库的数据,共删除一万多条数据。

(2) 进入 UNIX 系统压缩重组中文外部书目库。

(3) 将压缩重组好的中文外部数据库备份、打包,并传输到备份机备用。

这样,万一中文外部书目库出问题,可以用非常短的时间将备用的打包数据恢复,不至于影响图书馆正常的工作。

3.3 采访业务库的维护

3.3.1 回溯建库工作中存在的问题

编目工作人员在做书的回溯时,为了工作量的统计,一般要设置统一的“回溯批号”,但是有的工作人员由于疏忽,忘记设置“批号”,或者所设置的批号和上次的不相同,必须更正。编目工作人员在发现出错的情况下,一般采用将新做的那条采编馆藏数据删除,然后重新做一条正确的。

编目工作人员在做书的回溯,不仅在采编馆藏库增加了一条记录,同时也在采访业务库增加了一条记录,即“到书验收”记录,其“回溯批号”也就是采访业务库的“订购批标识”,因此,删除了采编馆藏库的数据,还应必须删除采访业务库同时新增的数据,否则会遗留垃圾数据。比较好的方法是在“回溯建库”界面的下方,点击“批号”栏,直接在上面修改。这样不仅更改了“回溯批号”,同时也修改了采访业务库的“订购批标识”。另一种方法,可以进入“采访管理”→“系统维护”→“采访业务库维护”界面,提取该数据进行修改更正,不过一般的工作人员没有进入“采访业务库维护”的权限。

3.3.2 采访业务库的数据批处理

尽管从“回溯建库”界面或从“采访业务库维护”界面都可以进行批号修改,但是这种方法都只能逐条逐条地修改,因此,要修改数十条,甚至是数百条的数据就要花费很多的时间和精力。ILASII 提供 MARC 数据和定长数据批处理接口,可以通过“系统管理员”→“系统工具”→“定长库数据批处理”,对“采访业务库”的批号进行批修改。无论书目数据是否已经从采编书目库审校交送到中央书目库,都可以采用这样的方法来处理。

4 结束语

我馆自 2001 年自动化集成系统升级为 ILASII2.0 后,笔者在几年的系统维护工作中,对 ILASII 系统的使用积累了一定的经验。本文对这些经验进行总结,希望对同仁的工作有所帮助。

[参考文献]

- [1] 洪丹萍. 关于 ILASII 数据备份与恢复的探讨[J]. 现代图书情报技术, 2002, (2): 91-93.
- [2] 王大可. ILASII 的开发模式与技术特点[J]. 现代图书情报技术, 1999, (2): 10-13

[作者简介] 洪丹萍(1955.7-),女,馆员,发表论文数篇。